PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-268663
(43)Date of publication of application : 20.09.2002

(51)Int.Cl. G10L 13/08
A63H 11/00
G10L 13/00
G10L 13/06
// B25J 13/00

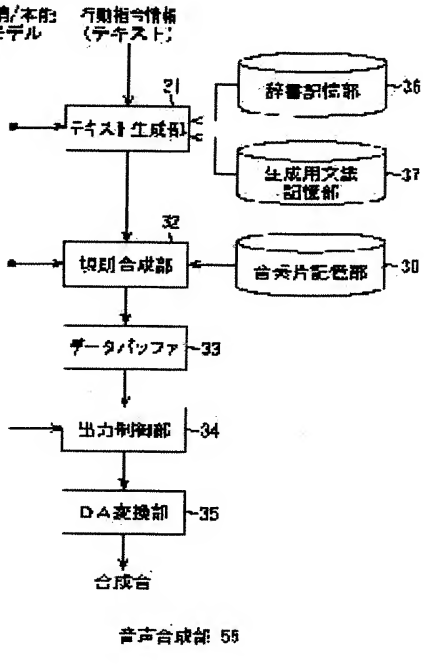
(21)Application number : 2001-065072
(22)Date of filing : 08.03.2001

(71)Applicant : SONY CORP
(72)Inventor : ASANO KOJI
KOBAYASHI KENICHIRO
YAMAZAKI NOBUHIDE
KARIYA SHINICHI
FUJITA YAEKO

(54) VOICE SYNTHESIZER, VOICE SYNTHESIS METHOD, PROGRAM AND RECORDING MEDIUM

(57)Abstract:
PROBLEM TO BE SOLVED: To actualize a pet robot, etc., having high interactivity.
SOLUTION: A text generation part 31 and a rule synthesis part 32 generate synthesized voice data corresponding to a text included in action command information according to the action command information and the data are stored in a data buffer 33. An output control part 34, on the other hand, controls the output of the synthesized voice data stored in the data buffer 33 according to the state of the feeling of the pet robot.

図5



音声合成部 55

LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号
特開2002－268663
（P2002－268663A）
(43)公開日　平成14年９月20日(2002.9.20)

(51)Int.Cl.[7]	識別記号	Ｆ Ｉ		テーマコード（参考）
Ｇ１０Ｌ　13/08		Ａ６３Ｈ　11/00	Ｚ	２Ｃ１５０
Ａ６３Ｈ　11/00		Ｂ２５Ｊ　13/00	Ｚ	３Ｃ００７
Ｇ１０Ｌ　13/00		Ｇ１０Ｌ　3/00	Ｈ	５Ｄ０４５
13/06			Ｑ	
// Ｂ２５Ｊ　13/00		5/04	Ｆ	

審査請求　未請求　請求項の数８　ＯＬ　（全 13 頁）

(21)出願番号　特願2001－65072(P2001－65072)
(22)出願日　平成13年３月８日(2001.3.8)

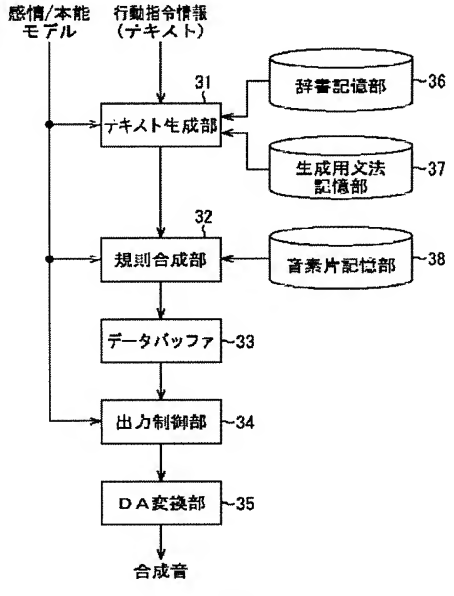
(71)出願人　000002185
ソニー株式会社
東京都品川区北品川６丁目７番35号
(72)発明者　浅野　康治
東京都品川区北品川６丁目７番35号　ソニー株式会社内
(72)発明者　小林　賢一郎
東京都品川区北品川６丁目７番35号　ソニー株式会社内
(74)代理人　100082131
弁理士　稲本　義雄

最終頁に続く

(54)【発明の名称】　音声合成装置および音声合成方法、並びにプログラムおよび記録媒体

(57)【要約】
【課題】　インタラクティブ性の高いペットロボット等を実現する。
【解決手段】　テキスト生成部３１および規則合成部３２では、行動指令情報にしたがい、その行動指令情報に含まれるテキストに対応する合成音データが生成され、データバッファ３３に記憶される。一方、出力制御部３４は、ペットロボットの感情の状態に基づき、データバッファ３３に記憶された合成音データの出力を制御する。

図5



音声合成部 55

【特許請求の範囲】
【請求項１】　情報処理装置の制御にしたがって、合成音を生成する音声合成装置であって、
前記情報処理装置の制御にしたがって、合成音を生成する音声合成手段と、
前記情報処理装置の内部状態に基づいて、前記合成音の出力を制御する出力制御手段とを備えることを特徴とする音声合成装置。
【請求項２】　前記出力制御手段は、前記情報処理装置の内部状態に基づいて、前記合成音の出力を停止することを特徴とする音声合成装置。
【請求項３】　前記音声合成手段は、前記出力制御手段が前記合成音の出力を停止した後、前記情報処理装置の内部状態に基づいて、前記合成音を再生成し、
前記出力制御手段は、前記音声合成手段において再生成された前記合成音を出力することを特徴とする請求項２に記載の音声合成装置。
【請求項４】　前記情報処理装置は、実際の、または仮想的なロボットであることを特徴とする請求項１に記載の音声合成装置。
【請求項５】　前記出力制御手段は、前記ロボットの感情または本能の状態に基づいて、前記合成音の出力を制御することを特徴とする請求項４に記載の音声合成装置。
【請求項６】　情報処理装置の制御にしたがって、合成音を生成する音声合成方法であって、
前記情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、
前記情報処理装置の内部状態に基づいて、前記合成音の出力を制御する出力制御ステップとを備えることを特徴とする音声合成方法。
【請求項７】　情報処理装置の制御にしたがって、合成音を生成する音声合成処理を、コンピュータに行わせる
{(　’　　/　　)P3(’　/’　)・・・)
・・・（１）
プログラムであって、
前記情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、
前記情報処理装置の内部状態に基づいて、前記合成音の出力を制御する出力制御ステップとを備えることを特徴とするプログラム。
【請求項８】　情報処理装置の制御にしたがって、合成音を生成する音声合成処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、
前記情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、
前記情報処理装置の内部状態に基づいて、前記合成音の出力を制御する出力制御ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、音声合成装置および音声合成方法、並びにプログラムおよび記録媒体に関し、特に、例えば、エンタテイメント用のロボットの感情等の状態に基づいて、合成音出力を制御するようにすることで、インタラクティブ性の高いロボットを実現すること等ができるようにする音声合成装置および音声合成方法、並びにプログラムおよび記録媒体に関する。
【０００２】
【従来の技術】従来の音声合成装置においては、テキスト、またはそのテキストを解析して得られる発音記号に基づいて、合成音が生成される。また、音声合成装置では、例えば、合成音の発話速度や、高さ、強さ、ポーズの長さ等が、音声合成装置に入力されるテキストや発音記号に挿入されたタグに基づいて制御される。
【０００３】ここで、タグが挿入された発音記号としては、例えば、次のようなものがある。
【０００４】
【０００５】（１）においては、発音記号をカタカナで表している。また、タグ’は、アクセントを表し、タグ{}および()は、フレーズの区切りを表す。さらに、タグP3のうちのPは、ポーズを表し、続く数字の3は、ポーズの長さを表す。また、タグ/は、アクセント句の区切りを表す。
【０００６】（１）の発音記号によれば、音声合成装置では、音韻が「これまで進められた研究は、大きな成果を・・・」という合成音が生成される。
【０００７】
【発明が解決しようとする課題】ところで、最近、例えば、ペット型のペットロボット等として、音声合成装置を搭載し、ユーザに話しかけたり、ユーザと会話（対話）を行うものが提案されている。
【０００８】さらに、ペットロボットとしては、感情の状態を表す感情モデルを取り入れ、その感情モデルが表す感情の状態に応じて、合成音の発話速度や、高さ、強さ、ポーズの長さ等を制御し、感情を表現した合成音を出力するものも提案されている。
【０００９】なお、音声に含まれる発話意図や感情と、発話速度や基本周波数等との関係については、例えば、前川、「音声によるパラ言語情報の伝達：言語学の立場から」、日本音響学会平成９年度秋季研究発表会講演論文集、pp.381-384（以下、文献１という）等に記載されている。
【００１０】ところで、例えば、上述のような音声合成装置を搭載したペットロボットにおいて、ユーザとのインタラクティブ性を向上させるためには、合成音を出力している最中に、感情の状態の変化に応じて、合成音の出力を停止したり、合成音の発話速度や高さ等を変更す

ることができるようにするのが望ましい。

【００１１】しかしながら、従来の音声合成装置では、上述したようなタグが挿入された発音記号列にしたがって、合成音が生成されるため、合成音が出力されている最中に、感情の状態が変化した場合に、リアルタイムで、合成音の出力を停止したりすること等が困難であった。

【００１２】本発明は、このような状況に鑑みてなされたものであり、ロボットの感情等の状態に応じて、合成音の出力を制御することにより、インタラクティブ性の高いロボット等を実現することができるようにするものである。

【００１３】

【課題を解決するための手段】本発明の音声合成装置は、情報処理装置の制御にしたがって、合成音を生成する音声合成手段と、情報処理装置の内部状態に基づいて、合成音の出力を制御する出力制御手段とを備えることを特徴とする。

【００１４】本発明の音声合成方法は、情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、情報処理装置の内部状態に基づいて、合成音の出力を制御する出力制御ステップとを備えることを特徴とする。

【００１５】本発明のプログラムは、情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、情報処理装置の内部状態に基づいて、合成音の出力を制御する出力制御ステップとを備えることを特徴とする。

【００１６】本発明の記録媒体は、情報処理装置の制御にしたがって、合成音を生成する音声合成ステップと、情報処理装置の内部状態に基づいて、合成音の出力を制御する出力制御ステップとを備えるプログラムが記録されていることを特徴とする。

【００１７】本発明の音声合成装置および音声合成方法、並びにプログラムにおいては、情報処理装置の制御にしたがって、合成音が生成される一方、情報処理装置の内部状態に基づいて、合成音の出力が制御される。

【００１８】

【発明の実施の形態】図１は、本発明を適用したロボットの一実施の形態の外観構成例を示しており、図２は、その電気的構成例を示している。

【００１９】本実施の形態では、ロボットは、例えば、犬等の四つ足の動物の形状のものとなっており、胴体部ユニット２の前後左右に、それぞれ脚部ユニット３Ａ，３Ｂ，３Ｃ，３Ｄが連結されるとともに、胴体部ユニット２の前端部と後端部に、それぞれ頭部ユニット４と尻尾部ユニット５が連結されることにより構成されている。

【００２０】尻尾部ユニット５は、胴体部ユニット２の上面に設けられたベース部５Ｂから、２自由度をもって湾曲または揺動自在に引き出されている。

【００２１】胴体部ユニット２には、ロボット全体の制御を行うコントローラ１０、ロボットの動力源となるバッテリ１１、並びにバッテリセンサ１２および熱センサ１３からなる内部センサ部１４などが収納されている。

【００２２】頭部ユニット４には、「耳」に相当するマイク（マイクロフォン）１５、「目」に相当するＣＣＤ(Charge Coupled Device)カメラ１６、触覚に相当するタッチセンサ１７、「口」に相当するスピーカ１８などが、それぞれ所定位置に配設されている。また、頭部ユニット４には、口の下顎に相当する下顎部４Ａが１自由度をもって可動に取り付けられており、この下顎部４Ａが動くことにより、ロボットの口の開閉動作が実現されるようになっている。

【００２３】脚部ユニット３Ａ乃至３Ｄそれぞれの関節部分や、脚部ユニット３Ａ乃至３Ｄそれぞれと胴体部ユニット２の連結部分、頭部ユニット４と胴体部ユニット２の連結部分、頭部ユニット４と下顎部４Ａの連結部分、並びに尻尾部ユニット５と胴体部ユニット２の連結部分などには、図２に示すように、それぞれアクチュエータ$3AA_1$乃至$3AA_K$、$3BA_1$乃至$3BA_K$、$3CA_1$乃至$3CA_K$、$3DA_1$乃至$3DA_K$、$4A_1$乃至$4A_L$、$5A_1$および$5A_2$が配設されている。

【００２４】頭部ユニット４におけるマイク１５は、ユーザからの発話を含む周囲の音声（音）を集音し、得られた音声信号を、コントローラ１０に送出する。ＣＣＤカメラ１６は、周囲の状況を撮像し、得られた画像信号を、コントローラ１０に送出する。

【００２５】タッチセンサ１７は、例えば、頭部ユニット４の上部に設けられており、ユーザからの「なでる」や「たたく」といった物理的な働きかけにより受けた圧力を検出し、その検出結果を圧力検出信号としてコントローラ１０に送出する。

【００２６】胴体部ユニット２におけるバッテリセンサ１２は、バッテリ１１の残量を検出し、その検出結果を、バッテリ残量検出信号としてコントローラ１０に送出する。熱センサ１３は、ロボット内部の熱を検出し、その検出結果を、熱検出信号としてコントローラ１０に送出する。

【００２７】コントローラ１０は、ＣＰＵ(Central Processing Unit)１０Ａやメモリ１０Ｂ等を内蔵しており、ＣＰＵ１０Ａにおいて、メモリ１０Ｂに記憶された制御プログラムが実行されることにより、各種の処理を行う。

【００２８】即ち、コントローラ１０は、マイク１５や、ＣＣＤカメラ１６、タッチセンサ１７、バッテリセンサ１２、熱センサ１３から与えられる音声信号、画像信号、圧力検出信号、バッテリ残量検出信号、熱検出信号に基づいて、周囲の状況や、ユーザからの指令、ユーザからの働きかけなどの有無を判断する。

【００２９】さらに、コントローラ１０は、この判断結果等に基づいて、続く行動を決定し、その決定結果に基づいて、アクチュエータ３AA$_1$乃至３AA$_K$、３BA$_1$乃至３BA$_K$、３CA$_1$乃至３CA$_K$、３DA$_1$乃至３DA$_K$、４A$_1$乃至４A$_L$、５A$_1$、５A$_2$のうちの必要なものを駆動させる。これにより、頭部ユニット４を上下左右に振らせたり、下顎部４Aを開閉させる。さらには、尻尾部ユニット５を動かせたり、各脚部ユニット３A乃至３Dを駆動して、ロボットを歩行させるなどの行動を行わせる。

【００３０】また、コントローラ１０は、必要に応じて、合成音を生成し、スピーカ１８に供給して出力させたり、ロボットの「目」の位置に設けられた図示しないＬＥＤ（Light Emitting Diode）を点灯、消灯または点滅させる。

【００３１】以上のようにして、ロボットは、周囲の状況等に基づいて自律的に行動をとるようになっている。

【００３２】次に、図3は、図２のコントローラ１０の機能的構成例を示している。なお、図３に示す機能的構成は、ＣＰＵ１０Aが、メモリ１０Bに記憶された制御プログラムを実行することで実現されるようになっている。

【００３３】コントローラ１０は、特定の外部状態を認識するセンサ入力処理部５０、センサ入力処理部５０の認識結果を累積して、感情や、本能、成長の状態を表現するモデル記憶部５１、センサ入力処理部５０の認識結果等に基づいて、続く行動を決定する行動決定機構部５２、行動決定機構部５２の決定結果に基づいて、実際にロボットに行動を起こさせる姿勢遷移機構部５３、各アクチュエータ３AA$_1$乃至５A$_1$および５A$_2$を駆動制御する制御機構部５４、並びに合成音を生成する音声合成部５５から構成されている。

【００３４】センサ入力処理部５０は、マイク１５や、ＣＣＤカメラ１６、タッチセンサ１７等から与えられる音声信号、画像信号、圧力検出信号等に基づいて、特定の外部状態や、ユーザからの特定の働きかけ、ユーザからの指示等を認識し、その認識結果を表す状態認識情報を、モデル記憶部５１および行動決定機構部５２に通知する。

【００３５】即ち、センサ入力処理部５０は、音声認識部５０Aを有しており、音声認識部５０Aは、マイク１５から与えられる音声信号について音声認識を行う。そして、音声認識部５０Aは、その音声認識結果としての、例えば、「歩け」、「伏せ」、「ボールを追いかけろ」等の指令その他を、状態認識情報として、モデル記憶部５１および行動決定機構部５２に通知する。

【００３６】また、センサ入力処理部５０は、画像認識部５０Bを有しており、画像認識部５０Bは、ＣＣＤカメラ１６から与えられる画像信号を用いて、画像認識処理を行う。そして、画像認識部５０Bは、その処理の結果、例えば、「赤い丸いもの」や、「地面に対して垂直なかつ所定高さ以上の平面」等を検出したときには、「ボールがある」や、「壁がある」等の画像認識結果を、状態認識情報として、モデル記憶部５１および行動決定機構部５２に通知する。

【００３７】さらに、センサ入力処理部５０は、圧力処理部５０Cを有しており、圧力処理部５０Cは、タッチセンサ１７から与えられる圧力検出信号を処理する。そして、圧力処理部５０Cは、その処理の結果、所定の閾値以上で、かつ短時間の圧力を検出したときには、「たたかれた（しかられた）」と認識し、所定の閾値未満で、かつ長時間の圧力を検出したときには、「なでられた（ほめられた）」と認識して、その認識結果を、状態認識情報として、モデル記憶部５１および行動決定機構部５２に通知する。

【００３８】モデル記憶部５１は、ロボットの感情、本能、成長の状態を表現する感情モデル、本能モデル、成長モデルをそれぞれ記憶、管理している。

【００３９】ここで、感情モデルは、例えば、「うれしさ」、「悲しさ」、「怒り」、「楽しさ」等の感情の状態（度合い）を、所定の範囲（例えば、－１．０乃至１．０等）の値によってそれぞれ表し、センサ入力処理部５０からの状態認識情報や時間経過等に基づいて、その値を変化させる。本能モデルは、例えば、「食欲」、「睡眠欲」、「運動欲」等の本能による欲求の状態（度合い）を、所定の範囲の値によってそれぞれ表し、センサ入力処理部５０からの状態認識情報や時間経過等に基づいて、その値を変化させる。成長モデルは、例えば、「幼年期」、「青年期」、「熟年期」、「老年期」等の成長の状態（度合い）を、所定の範囲の値によってそれぞれ表し、センサ入力処理部５０からの状態認識情報や時間経過等に基づいて、その値を変化させる。

【００４０】モデル記憶部５１は、上述のようにして感情モデル、本能モデル、成長モデルの値で表される感情、本能、成長の状態を、状態情報として、行動決定機構部５２に送出する。

【００４１】なお、モデル記憶部５１には、センサ入力処理部５０から状態認識情報が供給される他、行動決定機構部５２から、ロボットの現在または過去の行動、具体的には、例えば、「長時間歩いた」などの行動の内容を示す行動情報が供給されるようになっており、モデル記憶部５１は、同一の状態認識情報が与えられても、行動情報が示すロボットの行動に応じて、異なる状態情報を生成するようになっている。

【００４２】即ち、例えば、ロボットが、ユーザに挨拶をし、ユーザに頭を撫でられた場合には、ユーザに挨拶をしたという行動情報と、頭を撫でられたという状態認識情報とが、モデル記憶部５１に与えられ、この場合、モデル記憶部５１では、「うれしさ」を表す感情モデルの値が増加される。

【００４３】一方、ロボットが、何らかの仕事を実行中に頭を撫でられた場合には、仕事を実行中であるという行動情報と、頭を撫でられたという状態認識情報とが、モデル記憶部５１に与えられ、この場合、モデル記憶部５１では、「うれしさ」を表す感情モデルの値は変化されない。

【００４４】このように、モデル記憶部５１は、状態認識情報だけでなく、現在または過去のロボットの行動を示す行動情報も参照しながら、感情モデルの値を設定する。これにより、例えば、何らかのタスクを実行中に、ユーザが、いたずらするつもりで頭を撫でたときに、「うれしさ」を表す感情モデルの値を増加させるような、不自然な感情の変化が生じることを回避することができる。

【００４５】なお、モデル記憶部５１は、本能モデルおよび成長モデルについても、感情モデルにおける場合と同様に、状態認識情報および行動情報の両方に基づいて、その値を増減させるようになっている。また、モデル記憶部５１は、感情モデル、本能モデル、成長モデルそれぞれの値を、他のモデルの値にも基づいて増減させるようになっている。

【００４６】行動決定機構部５２は、センサ入力処理部５０からの状態認識情報や、モデル記憶部５１からの状態情報、時間経過等に基づいて、次の行動を決定し、決定された行動の内容を、行動指令情報として、姿勢遷移機構部５３に送出する。

【００４７】即ち、行動決定機構部５２は、ロボットがとり得る行動をステート（状態）(state)に対応させた有限オートマンを、ロボットの行動を規定する行動モデルとして管理しており、この行動モデルとしての有限オートマンにおけるステートを、センサ入力処理部５０からの状態認識情報や、モデル記憶部５１における感情モデル、本能モデル、または成長モデルの値、時間経過等に基づいて遷移させ、遷移後のステートに対応する行動を、次にとるべき行動として決定する。

【００４８】ここで、行動決定機構部５２は、所定の取りが(trigger)があったことを検出すると、ステートを遷移させる。即ち、行動決定機構部５２は、例えば、現在のステートに対応する行動を実行している時間が所定時間に達したときや、特定の状態認識情報を受信したとき、モデル記憶部５１から供給される状態情報が示す感情や、本能、成長の状態の値が所定の閾値以下または以上になったとき等に、ステートを遷移させる。

【００４９】なお、行動決定機構部５２は、上述したように、センサ入力処理部５０からの状態認識情報だけでなく、モデル記憶部５１における感情モデルや、本能モデル、成長モデルの値等に基づいて、行動モデルにおけるステートを遷移させることから、同一の状態認識情報が入力されても、感情モデルや、本能モデル、成長モデルの値（状態情報）によっては、ステートの遷移先は異なるものとなる。

【００５０】その結果、行動決定機構部５２は、例えば、状態情報が、「怒っていない」こと、および「お腹がすいていない」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表しているときには、目の前に手のひらが差し出されたことに応じて、「お手」という行動をとらせる行動指令情報を生成し、これを、姿勢遷移機構部５３に送出する。

【００５１】また、行動決定機構部５２は、例えば、状態情報が、「怒っていない」こと、および「お腹がすいている」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表しているときには、目の前に手のひらが差し出されたことに応じて、「手のひらをぺろぺろなめる」ような行動を行わせるための行動指令情報を生成し、これを、姿勢遷移機構部５３に送出する。

【００５２】また、行動決定機構部５２は、例えば、状態情報が、「怒っている」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表しているときには、状態情報が、「お腹がすいている」ことを表していても、また、「お腹がすいていない」ことを表していても、「ぷいと横を向く」ような行動を行わせるための行動指令情報を生成し、これを、姿勢遷移機構部５３に送出する。

【００５３】なお、行動決定機構部５２には、モデル記憶部５１から供給される状態情報が示す感情や、本能、成長の状態に基づいて、遷移先のステートに対応する行動のパラメータとしての、例えば、歩行の速度や、手足を動かす際の動きの大きさおよび速度などを決定させることができ、この場合、それらのパラメータを含む行動指令情報が、姿勢遷移機構部５３に送出される。

【００５４】また、行動決定機構部５２では、上述したように、ロボットの頭部や手足等を動作させる行動指令情報の他、ロボットに発話を行わせる行動指令情報も生成される。ロボットに発話を行わせる行動指令情報は、音声合成部５５に供給されるようになっており、音声合成部５５に供給される行動指令情報には、音声合成部５５に生成させる合成音に対応するテキスト等が含まれる。そして、音声合成部５５は、行動決定部５２から行動指令情報を受信すると、その行動指令情報に含まれるテキストに基づき、合成音を生成し、スピーカ１８に供給して出力させる。これにより、スピーカ１８からは、例えば、ロボットの鳴き声、さらには、「お腹がすいた」等のユーザへの各種の要求、「何？」等のユーザの呼びかけに対する応答その他の音声出力が行われる。ここで、音声合成部５５には、モデル記憶部５１から状態情報も供給されるようになっており、音声合成部５５は、この状態情報が示す感情の状態に基づいて韻律等を制御した合成音を生成することが可能となっている。な

お、音声合成部５５では、感情の他、本能や本能の状態に基づいて韻律等を制御した合成音を生成することも可能である。また、行動決定機構部５２は、合成音を出力する場合には、下顎部４Ａを開閉させる行動指令情報を、必要に応じて生成し、姿勢遷移機構部５３に出力する。この場合、合成音の出力に同期して、下顎部４Ａが開閉し、ユーザに、ロボットがしゃべっているかのような印象を与えることができる。

【００５５】姿勢遷移機構部５３は、行動決定機構部５２から供給される行動指令情報に基づいて、ロボットの姿勢を、現在の姿勢から次の姿勢に遷移させるための姿勢遷移情報を生成し、これを制御機構部５４に送出する。

【００５６】ここで、現在の姿勢から次に遷移可能な姿勢は、例えば、胴体や手や足の形状、重さ、各部の結合状態のようなロボットの物理的形状と、関節が曲がる方向や角度のようなアクチュエータ３ＡＡ$_1$乃至５Ａ$_1$および５Ａ$_2$の機構とによって決定される。

【００５７】また、次の姿勢としては、現在の姿勢から直接遷移可能な姿勢と、直接には遷移できない姿勢とがある。例えば、４本足のロボットは、手足を大きく投げ出して寝転んでいる状態から、伏せた状態へ直接遷移することはできるが、立った状態へ直接遷移することはできず、一旦、手足を胴体近くに引き寄せて伏せた姿勢になり、それから立ち上がるという２段階の動作が必要である。また、安全に実行できない姿勢も存在する。例えば、４本足のロボットは、その４本足で立っている姿勢から、両前足を挙げてバンザイをしようとすると、簡単に転倒してしまう。

【００５８】このため、姿勢遷移機構部５３は、直接遷移可能な姿勢をあらかじめ登録しておき、行動決定機構部５２から供給される行動指令情報が、直接遷移可能な姿勢を示す場合には、その行動指令情報を、そのまま姿勢遷移情報として、制御機構部５４に送出する。一方、行動指令情報が、直接遷移不可能な姿勢を示す場合には、姿勢遷移機構部５３は、遷移可能な他の姿勢に一旦遷移した後に、目的の姿勢まで遷移させるような姿勢遷移情報を生成し、制御機構部５４に送出する。これによりロボットが、遷移不可能な姿勢を無理に実行しようとする事態や、転倒するような事態を回避することができるようになっている。

【００５９】制御機構部５４は、姿勢遷移機構部５３からの姿勢遷移情報にしたがって、アクチュエータ３ＡＡ$_1$乃至５Ａ$_1$および５Ａ$_2$を駆動するための制御信号を生成し、これを、アクチュエータ３ＡＡ$_1$乃至５Ａ$_1$および５Ａ$_2$に送出する。これにより、アクチュエータ３ＡＡ$_1$乃至５Ａ$_1$および５Ａ$_2$は、制御信号にしたがって駆動し、ロボットは、自律的に行動を起こす。

【００６０】次に、図４は、図３の音声認識部５０Ａの構成例を示している。

【００６１】マイク１５からの音声信号は、ＡＤ(Analog Digital)変換部２１に供給される。ＡＤ変換部２１では、マイク１５からのアナログ信号である音声信号がサンプリング、量子化され、ディジタル信号である音声データにＡ／Ｄ変換される。この音声データは、特徴抽出部２２および音声区間検出部２７に供給される。

【００６２】特徴抽出部２２は、そこに入力される音声データについて、適当なフレームごとに、例えば、ＭＦＣＣ(Mel Frequency Cepstrum Coefficient)分析を行い、その分析の結果得られるＭＦＣＣを、特徴パラメータ（特徴ベクトル）として、マッチング部２３に出力する。なお、特徴抽出部２２では、その他、例えば、線形予測係数、ケプストラム係数、線スペクトル対、所定の周波数帯域ごとのパワー（フィルタバンクの出力）等を、特徴パラメータとして抽出することが可能である。

【００６３】マッチング部２３は、特徴抽出部２２からの特徴パラメータを用いて、音響モデル記憶部２４、辞書記憶部２５、および文法記憶部２６を必要に応じて参照しながら、マイク１５に入力された音声（入力音声）を、例えば、連続分布ＨＭＭ(Hidden Markov Model)法に基づいて音声認識する。

【００６４】即ち、音響モデル記憶部２４は、音声認識する音声の言語における個々の音素や音節などの音響的な特徴を表す音響モデルを記憶している。ここでは、連続分布ＨＭＭ法に基づいて音声認識を行うので、音響モデルとしては、ＨＭＭ(Hidden Markov Model)が用いられる。辞書記憶部２５は、認識対象の各単語について、その発音に関する情報（音韻情報）が記述された単語辞書を記憶している。文法記憶部２６は、辞書記憶部２５の単語辞書に登録されている各単語が、どのように連鎖する（つながる）かを記述した文法規則を記憶している。ここで、文法規則としては、例えば、文脈自由文法（ＣＦＧ）や、統計的な単語連鎖確率（Ｎ－ｇｒａｍ）などに基づく規則を用いることができる。

【００６５】マッチング部２３は、辞書記憶部２５の単語辞書を参照することにより、音響モデル記憶部２４に記憶されている音響モデルを接続することで、単語の音響モデル（単語モデル）を構成する。さらに、マッチング部２３は、幾つかの単語モデルを、文法記憶部２６に記憶された文法規則を参照することにより接続し、そのようにして接続された単語モデルを用いて、特徴パラメータに基づき、連続分布ＨＭＭ法によって、マイク１５に入力された音声を認識する。即ち、マッチング部２３は、特徴抽出部２２が出力する時系列の特徴パラメータが観測されるスコア（尤度）が最も高い単語モデルの系列を検出し、その単語モデルの系列に対応する単語列の音韻情報（読み）を、音声の認識結果として出力する。

【００６６】より具体的には、マッチング部２３は、接続された単語モデルに対応する単語列について、各特徴パラメータの出現確率（出力確率）を累積し、その累積

値をスコアとして、そのスコアを最も高くする単語列の音韻情報を、音声認識結果として出力する。

【００６７】以上のようにして出力される、マイク１５に入力された音声の認識結果は、状態認識情報として、モデル記憶部５１および行動決定機構部５２に出力される。

【００６８】なお、音声区間検出部２７は、ＡＤ変換部２１からの音声データについて、特徴抽出部２２がＭＦＣＣ分析を行うのと同様のフレームごとに、例えば、パワーを算出している。さらに、音声区間検出部２７は、各フレームのパワーを、所定の閾値と比較し、その閾値以上のパワーを有するフレームで構成される区間を、ユーザの音声が入力されている音声区間として検出する。そして、音声区間検出部２７は、検出した音声区間を、特徴抽出部２２とマッチング部２３に供給しており、特徴抽出部２２とマッチング部２３は、音声区間のみを対象に処理を行う。

【００６９】次に、図5は、図3の音声合成部５５の構成例を示している。

【００７０】テキスト生成部３１には、行動決定機構部５２が出力する、音声合成の対象とするテキストを含む行動指令情報が供給されるようになっており、テキスト生成部３１は、辞書記憶部３６や生成用文法記憶部３７を参照しながら、その行動指令情報に含まれるテキストを解析する。

【００７１】即ち、辞書記憶部３６には、各単語の品詞情報や、読み、アクセント等の情報が記述された単語辞書が記憶されており、また、生成用文法記憶部３７には、辞書記憶部３６の単語辞書に記述された単語について、単語連鎖に関する制約等の文法規則が記憶されている。そして、テキスト生成部３１は、この単語辞書および文法規則に基づいて、そこに入力されるテキストの形態素解析や構文解析等の解析を行い、後段の規則合成部３２で行われる規則音声合成に必要な情報を抽出する。ここで、規則音声合成に必要な情報としては、例えば、ポーズの位置や、アクセントおよびイントネーションを制御するための情報その他の韻律情報や、各単語の発音等の音韻情報などがある。

【００７２】テキスト生成部３１で得られた情報は、規則合成部３２に供給され、規則合成部３２は、音素片記憶部３８を用いて、テキスト生成部３１に入力されたテキストに対応する合成音の音声データ（ディジタルデータ）を生成する。

【００７３】即ち、音素片記憶部３８には、例えば、ＣＶ(Consonant, Vowel)や、ＶＣＶ、ＣＶＣ等の形で音素片データが記憶されており、規則合成部３２は、テキスト生成部３１からの情報に基づいて、必要な音素片データを接続し、さらに、ポーズ、アクセント、イントネーション等を適切に付加することで、テキスト生成部３１に入力されたテキストに対応する合成音データを生成する。

【００７４】この音声データは、データバッファ３３に供給される。データバッファ３３は、規則合成部３２から供給される合成音データを記憶する。

【００７５】出力制御部３４は、定期的または不定期に、モデル記憶部５１（図3）に記憶された感情モデル等をチェックし、その感情モデル等に基づいて、データバッファ３３に記憶された合成音データの出力を制御する。

【００７６】即ち、出力制御部３４は、感情モデルの値（感情モデル値）が、ある条件を満たすとき、データバッファ３３に記憶された合成音データを読み出し、ＤＡ（Digital Analogue）変換部３５に供給する。この場合、ＤＡ変換部３５は、ディジタル信号としての合成音データを、アナログ信号としての音声信号にＤ／Ａ変換する。この音声信号は、スピーカ１８に供給され、これにより、テキスト生成部３１に入力されたテキストに対応する合成音が出力される。

【００７７】また、出力制御部３４は、感情モデル値が、他の条件を満たすとき、データバッファ３３に記憶された合成音データの読み出しを停止する。この場合、スピーカ１８からの合成音の出力は停止する。

【００７８】また、出力制御部３４は、感情モデル値が、さらに他の条件を満たすとき、データバッファ３３に記憶された合成音データの読み出しを停止し、その後、データバッファ３３に記憶された合成音データの読み出しを再開する。この場合、スピーカ１８からの合成音の出力は、一旦停止され、その後、再開される。

【００７９】なお、テキスト生成部３１および規則合成部３２も、出力制御部３４と同様に、モデル記憶部５１（図3）に記憶された感情モデルの値（感情モデル値）や本能モデルの値（本能モデル値）をチェックするようになっており、この感情モデル値や本能モデル値を考慮して処理を行うようになっている。

【００８０】次に、図6のフローチャートを参照して、図5の音声合成部５５による音声合成処理について説明する。

【００８１】行動決定機構部５２が、音声合成の対象とするテキストを含む行動指令情報を、音声合成部５５に出力すると、テキスト生成部３１は、ステップＳ１において、その行動指令情報を受信し、ステップＳ２に進む。ステップＳ２では、テキスト生成部３１および規則合成部３２において、モデル記憶部５１を参照することで、感情モデル値や本能モデル値が認識（チェック）され、ステップＳ３に進む。

【００８２】ステップＳ３では、テキスト生成部３１において、行動決定機構部５２からの行動指令情報に含まれるテキストから、実際に合成音として出力するテキスト（以下、適宜、発話テキストという）を生成する際に用いる語彙（発話語彙）が、感情モデル値や本能モデル

値に基づいて設定され、ステップＳ４に進む。ステップＳ４では、テキスト生成部３１において、ステップＳ３で設定された発話語彙を用いて、行動指令情報に含まれるテキストに対応する発話テキストが生成される。

【００８３】即ち、行動決定機構部５２からの行動指令情報に含まれるテキストは、例えば、標準的な感情および本能の状態における発話を前提としたものとなっており、ステップＳ４では、そのテキストが、ロボットの感情や本能の状態を考慮して修正され、これにより、発話テキストが生成される。

【００８４】具体的には、例えば、行動指令情報に含まれるテキストが、「何ですか」である場合において、ロボットの感情の状態が「怒っている」ことを表しているときには、その怒りを表現する「何だよ！」が、発話テキストとして生成される。あるいは、また、例えば、行動指令情報に含まれるテキストが、「やめて下さい」である場合において、ロボットの感情の状態が「怒っている」ことを表しているときには、その怒りを表現する「やめろ！」が、発話テキストとして生成される。

【００８５】そして、ステップＳ５に進み、テキスト生成部３１は、発話テキストを対象に、形態素解析や構文解析等のテキスト解析を行い、その発話テキストについて規則音声合成を行うのに必要な情報としての、ピッチ周波数や、パワー、継続時間長等の韻律情報を生成する。さらに、テキスト生成部３１は、発話テキストを構成する各単語の発音等の音韻情報も生成する。ここで、ステップＳ５では、発話テキストの韻律情報として、標準的な韻律情報が生成される。

【００８６】その後、テキスト生成部３１は、ステップＳ６において、ステップＳ５で設定した発話テキストの韻律情報を、ロボットの感情や本能の状態に基づいて修正し、これにより、発話テキストが合成音で出力されるときの感情表現が高められる。

【００８７】テキスト生成部３１で得られた発話テキストの音韻情報および韻律情報は、規則合成部３２に供給され、規則合成部３２では、ステップＳ７において、その音韻情報および韻律情報にしたがい、規則音声合成が行われることにより、発話テキストの合成音のディジタルデータ（合成音データ）が生成される。ここで、規則合成部３２でも、規則音声合成の際、感情モデル値や本能モデル値に基づいて、ロボットの感情や本能の状態を適切に表現するように、合成音のポーズの位置や、アクセントの位置、イントネーション等の韻律が変更される。

【００８８】規則合成部３２で得られた合成音データは、ステップＳ８において、データバッファ３３に供給され、データバッファ３３は、規則合成部３２からの合成音データを記憶する。

【００８９】そして、ステップＳ９に進み、出力制御部３４は、モデル記憶部５１に記憶された感情モデル値や本能モデル値をチェックし、ステップＳ１０に進む。ステップＳ１０では、出力制御部３４は、直前のステップＳ９においてチェックした感情モデル値や本能モデル値に基づき、合成音の出力を中断（停止）するかどうかを判定する。

【００９０】ステップＳ１０において、合成音の出力を中断しないと判定された場合、ステップＳ１１に進み、出力制御部３４は、データバッファ３３から所定量（例えば、１秒分）の合成音データを読み出し、ＤＡ変換部３５に供給する。ＤＡ変換部３５では、ステップＳ１２において、出力制御部３４からの合成音データがＤ／Ａ変換され、スピーカ１８に供給されて出力される。

【００９１】その後、ステップＳ１３に進み、出力制御部３４は、データバッファ３３に合成音データが記憶されていないかどうか、即ち、データバッファ３３が空かどうかを判定する。ステップＳ１３において、データバッファ３３が空でなく、まだ、合成音データが記憶されていると判定された場合、ステップＳ９に戻り、以下、同様の処理が繰り返される。

【００９２】従って、この場合は、規則合成部３２で生成された合成音が出力され続ける。

【００９３】また、ステップＳ１３において、データバッファ３３が空であると判定された場合、即ち、データバッファ３３に記憶された合成音データの出力が完了した場合、処理を終了する。

【００９４】一方、ステップＳ１０において、合成音の出力を中断すると判定された場合、ステップＳ１４に進み、出力制御部３４は、データバッファ３３をクリアし、ステップＳ１５に進む。ステップＳ１５では、出力制御部３４は、必要に応じて、行動決定機構部５２に対して、音声合成の対象とするテキストを含む行動指令情報の再出力を要求し、処理を終了する。

【００９５】従って、この場合は、合成音の出力が途中で停止される。そして、行動指令情報の再出力の要求が行われた場合には、行動決定機構部５２からの行動指令情報の再出力を待って、ステップＳ１からの処理が行われることにより、合成音の出力が、最初から再開される。

【００９６】以上のような音声合成処理において、ステップＳ１０における、合成音の出力を中断するかどうかの判定は、感情モデル値のうちの、例えば、「怒り」を表すものに基づいて行うことができる。即ち、「怒り」の度合いが高い場合に、合成音の出力を停止するようにすることができる。

【００９７】この場合、合成音が出力されている最中に、ユーザが、ペットロボットを叩く等して、「怒り」の度合いが高くなると、ペットロボットは、即座に合成音の出力を停止する。従って、この場合、ペットロボットが突然黙った状態となることによって、ユーザに対して、ペットロボットが怒った状態にあることを印象づけ

ることができる。

【００９８】さらに、「怒り」の度合いが高いが、それほどでもない場合には、出力制御部３４において、データバッファ３３をクリアした後、行動決定機構部５２に対して、行動指令情報の再出力を要求するようにすることができる。

【００９９】この場合、「怒り」の度合いが高くなっていることから、上述したように、テキスト生成部３１では、「怒り」を表現する発話テキストが生成され、さらに、規則合成部３２では、「怒り」を表現する韻律が付された合成音データが生成される。従って、この場合、スピーカ１８からは、「怒り」を表す合成音が出力され、その結果、ユーザには、ペットロボットが、いわば怒った口調で言い直しを行ったかのような印象を与えることができる。

【０１００】以上のように、ペットロボットの感情の状態に基づいて、合成音の出力を制御するようにしたので、インタラクティブ性の高いペットロボットを実現することができる。

【０１０１】なお、例えば、ユーザがシステムからの音声出力中に発話を行った場合に、いわゆるバージインに対処する必要から、システムの音声出力を中断する音声対話システムが研究されているが、これは、ユーザによる音声入力を遮らないようにするためであり、感情等のシステムの内部状態に基づいて、システムの音声出力を中断するものではない。従って、このような音声対話システムによれば、ユーザによる音声入力を妨げることを防止することはできるが、本実施の形態におけるペットロボットのように、ユーザとの間のインタラクティブ性を向上させることはできない。

【０１０２】以上、本発明を、エンターテイメント用のロボット（疑似ペットとしてのロボット）に適用した場合について説明したが、本発明は、これに限らず、例えば、音声合成装置を搭載した対話システムその他に広く適用することが可能である。また、本発明は、現実世界のロボットだけでなく、例えば、液晶ディスプレイ等の表示装置に表示される仮想的なロボットにも適用可能である。

【０１０３】なお、本実施の形態においては、上述した一連の処理を、ＣＰＵ１０Ａにプログラムを実行させることにより行うようにしたが、一連の処理は、それ専用のハードウェアによって行うことも可能である。

【０１０４】ここで、プログラムは、あらかじめメモリ１０Ｂ（図２）に記憶させておく他、フロッピーディスク、CD-ROM(Compact Disc Read Only Memory), MO(Magnetooptical)ディスク, DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体に、一時的あるいは永続的に格納（記録）しておくことができる。そして、このようなリムーバブル記録媒体を、いわゆるパッケージソフトウエアとして提供し、ロボット（メモリ１０Ｂ）にインストールするようにすることができる。

【０１０５】また、プログラムは、ダウンロードサイトから、ディジタル衛星放送用の人工衛星を介して、無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、有線で転送し、メモリ１０Ｂにインストールすることができる。

【０１０６】この場合、プログラムがバージョンアップされたとき等に、そのバージョンアップされたプログラムを、メモリ１０Ｂに、容易にインストールすることができる。

【０１０７】なお、本明細書において、ＣＰＵ１０Ａに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。

【０１０８】また、プログラムは、１のＣＰＵにより処理されるものであっても良いし、複数のＣＰＵによって分散処理されるものであっても良い。

【０１０９】次に、図５の音声合成装置５５は、専用のハードウェアにより実現することもできるし、ソフトウェアにより実現することもできる。音声合成装置５５をソフトウェアによって実現する場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

【０１１０】そこで、図7は、音声合成装置５５を実現するためのプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

【０１１１】プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク１０５やＲＯＭ１０３に予め記録しておくことができる。

【０１１２】あるいはまた、プログラムは、フロッピー（登録商標）ディスク、CD-ROM, MOディスク, DVD、磁気ディスク、半導体メモリなどのリムーバブル記録媒体１１１に、一時的あるいは永続的に格納（記録）しておくことができる。このようなリムーバブル記録媒体１１１は、いわゆるパッケージソフトウエアとして提供することができる。

【０１１３】なお、プログラムは、上述したようなリムーバブル記録媒体１１１からコンピュータにインストールする他、ダウンロードサイトから、ディジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを、通信部１０８で受信し、内蔵するハードディスク１０５にインストールすることができる。

【０１１４】コンピュータは、CPU(Central Processing Unit)１０２を内蔵している。CPU１０２には、バス１

０１を介して、入出力インタフェース１１０が接続されており、CPU１０２は、入出力インタフェース１１０を介して、ユーザによって、キーボードや、マウス、マイク等で構成される入力部１０７が操作等されることにより指令が入力されると、それにしたがって、ROM(Read Only Memory)１０３に格納されているプログラムを実行する。あるいは、また、CPU１０２は、ハードディスク１０５に格納されているプログラム、衛星若しくはネットワークから転送され、通信部１０８で受信されてハードディスク１０５にインストールされたプログラム、またはドライブ１０９に装着されたリムーバブル記録媒体１１１から読み出されてハードディスク１０５にインストールされたプログラムを、RAM(Random Access Memory)１０４にロードして実行する。これにより、CPU１０２は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、CPU１０２は、その処理結果を、必要に応じて、例えば、入出力インタフェース１１０を介して、LCD(Liquid CryStal Display)やスピーカ等で構成される出力部１０６から出力、あるいは、通信部１０８から送信、さらには、ハードディスク１０５に記録等させる。

【０１１５】なお、本実施の形態では、行動決定機構部５２が生成するテキストから合成音を生成するようにしたが、本発明は、あからじめ用意されたテキストから合成音を生成する場合にも適用可能である。さらに、本発明は、あらかじめ録音してある音声データを編集して、目的とする合成音を生成する場合にも適用可能である。

【０１１６】また、本実施の形態では、ペットロボットの感情の状態に基づいて、合成音の出力を制御するようにしたが、合成音の出力は、その他、例えば、本能や成長その他のペットロボットの内部状態に基づいて制御することが可能である。

【０１１７】

【発明の効果】以上の如く、本発明の音声合成装置および音声合成方法、並びにプログラムによれば、情報処理装置の制御にしたがって、合成音が生成される一方、情報処理装置の内部状態に基づいて、合成音の出力が制御される。従って、インタラクティブ性の高い合成音の出力を行うことが可能となる。

【図面の簡単な説明】

【図１】本発明を適用したロボットの一実施の形態の外観構成例を示す斜視図である。

【図２】ロボットの内部構成例を示すブロック図である。

【図３】コントローラ１０の機能的構成例を示すブロック図である。

【図４】音声認識部５０Aの構成例を示すブロック図である。

【図５】音声合成部５５の構成例を示すブロック図である。

【図６】音声合成部５５による音声合成処理を説明するフローチャートである。

【図７】本発明を適用したコンピュータの一実施の形態の構成例を示すブロック図である。

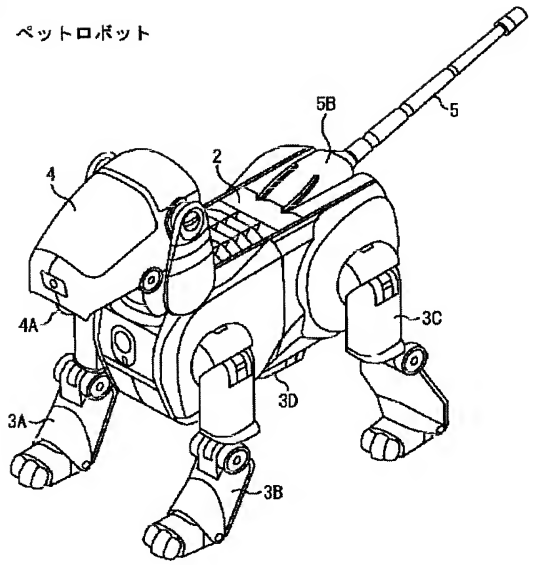
【符号の説明】

１　頭部ユニット，　４A　下顎部，　１０　コントローラ，　１０A　CPU，　１０B　メモリ，　１５　マイク，　１６　CCDカメラ，　１７　タッチセンサ，　１８　スピーカ，　２１　AD変換部，　２２　特徴抽出部，　２３　マッチング部，　２４　音響モデル記憶部，　２５　辞書記憶部，　２６文法記憶部，　２７　音声区間検出部，　３１　テキスト生成部，　３２　規則合成部，　３３　データバッファ，　３４　出力制御部，　３５　DA変換部，３６　辞書記憶部，　３７　生成用文法記憶部，　３８　音素片記憶部，　５０　センサ入力処理部，　５０A　音声認識部，　５０B　画像認識部，　５０C　圧力処理部，　５１　モデル記憶部，　５２　行動決定機構部，　５３　姿勢遷移機構部，　５４　制御機構部，　５５　音声合成部，　１０１　バス，１０２　CPU，　１０３　ROM，　１０４　RAM，　１０５　ハードディスク，１０６　出力部，　１０７　入力部，　１０８　通信部，　１０９　ドライブ，１１０　入出力インタフェース，　１１１　リムーバブル記録媒体

【図1】

図1

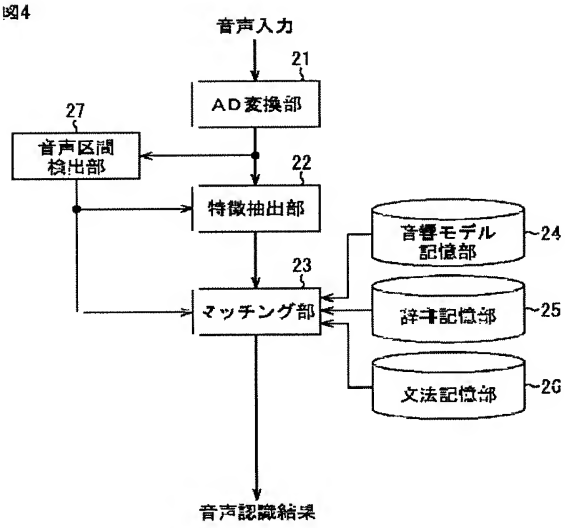
ペットロボット



ペットロボットの外観構成

【図4】

図4



音声認識部 50A

【図2】

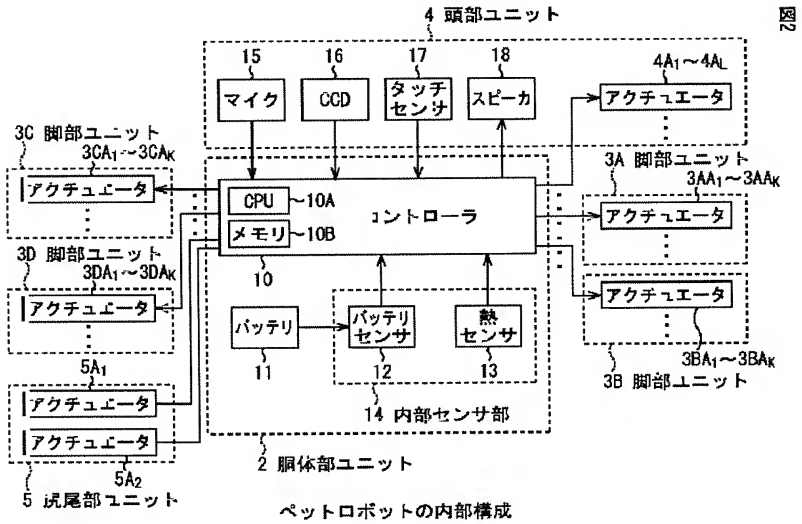


図2

ペットロボットの内部構成

【図3】

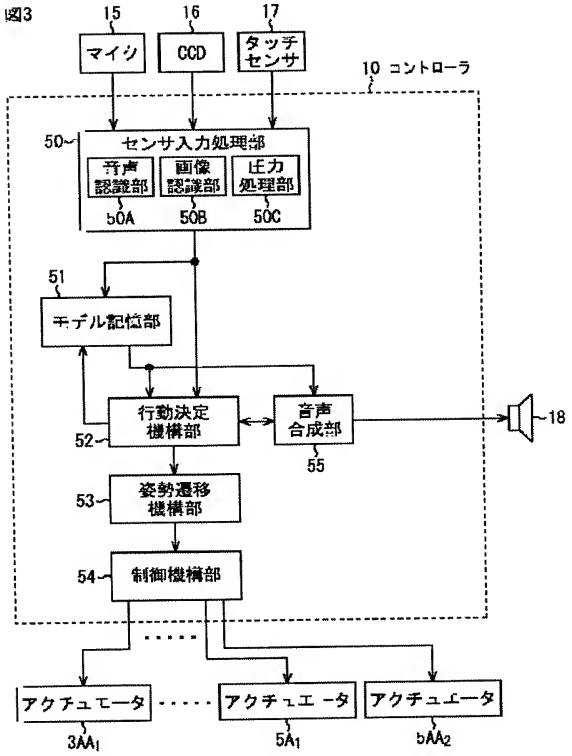


図3
15
16
17
マイク
CCD
タッチセンサ
10 コントローラ
50
センサ入力処理部
音声認識部
画像認識部
圧力処理部
50A
50B
50C
51
モデル記憶部
行動決定機構部
52
音声合成部
55
18
姿勢遷移機構部
53
制御機構部
54
アクチュエータ
アクチュエータ
アクチュエータ
3AA1
5A1
5AA2

【図5】

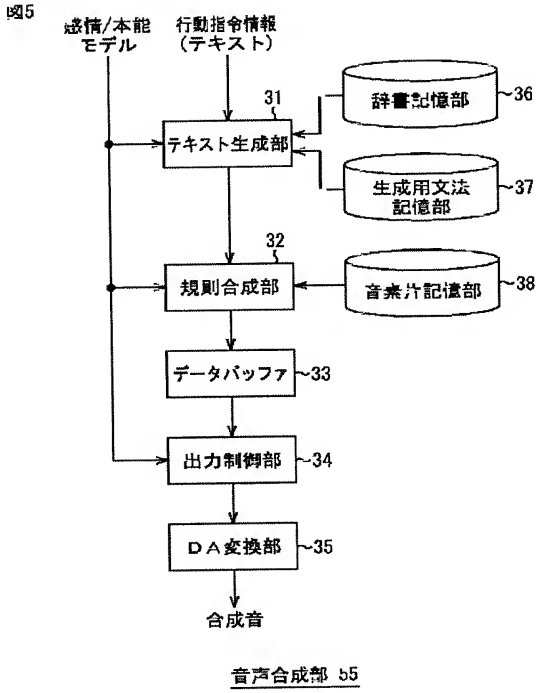


図5
感情/本能モデル
行動指令情報(テキスト)
31
テキスト生成部
辞書記憶部
36
生成用文法記憶部
37
32
規則合成部
音素片記憶部
38
データバッファ
33
出力制御部
34
DA変換部
35
合成音
音声合成部 55

【図7】

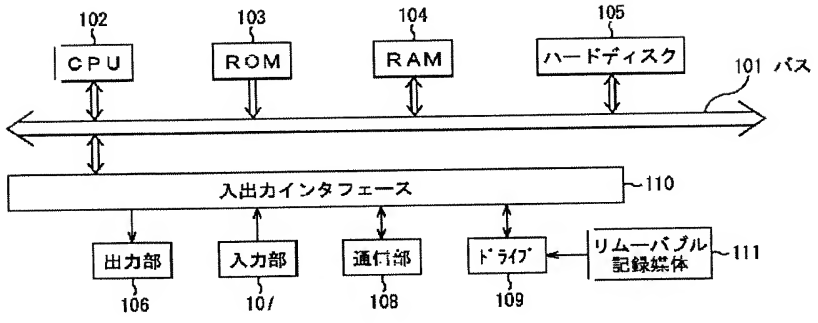
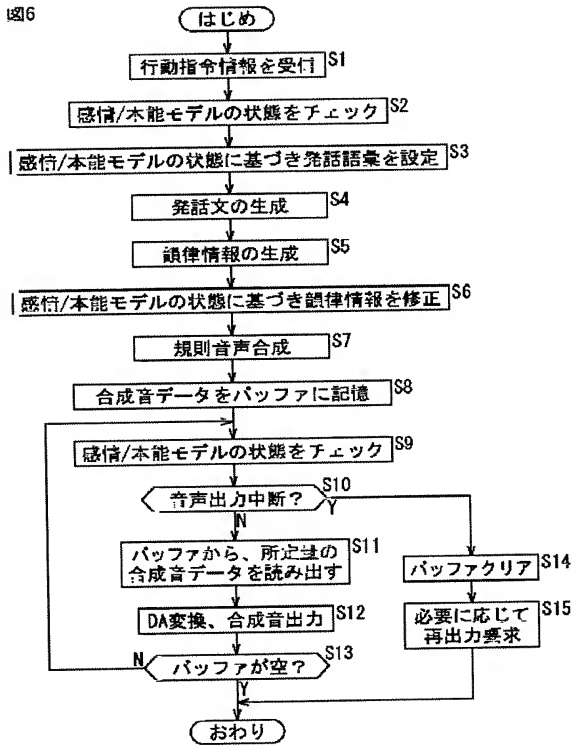


図7
102
103
104
105
CPU
ROM
RAM
ハードディスク
101 バス
入出力インタフェース
110
出力部
入力部
通信部
ドライブ
リムーバブル記録媒体
111
106
107
108
109
コンピュータ

【図６】



フロントページの続き

(72)発明者　山崎　信英
東京都品川区北品川６丁目７番35号　ソニー株式会社内
(72)発明者　狩谷　真一
東京都品川区北品川６丁目７番35号　ソニー株式会社内
(72)発明者　藤田　八重子
東京都品川区北品川６丁目７番35号　ソニー株式会社内

Ｆターム(参考)　2C150 CA01 CA02 CA04 DA05 DA24
DA25 DA26 DA27 DA28 DF03
DF04 DF33 ED42 ED52 EF03
EF07 EF09 EF13 EF16 EF23
EF29 EF34 EF36
3C007 AS36 CS08 KS10 MT14 WA04
WA14 WB16 WB28 WC30
5D045 AA08 AA09 AB11